Three-Dimensional Solution Structure of PsaE from the Cyanobacterium *Synechococcus* sp. Strain PCC 7002, a Photosystem I Protein That Shows Structural Homology with SH3 Domains[†,‡]

Christopher J. Falzone,[*,§] Yung-Hsiang Kao,[§] Jindong Zhao,[∥] Donald A. Bryant,[*,∥] and Juliette T. J. Lecomte[*,§]

Department of Chemistry, Department of Biochemistry and Molecular Biology, and Center for Biomolecular Structure and Function, The Pennsylvania State University, University Park, Pennsylvania 16802

Received January 5, 1994; Revised Manuscript Received March 9, 1994[⊗]

ABSTRACT: PsaE is a 69 amino acid polypeptide from photosystem I present on the stromal side of the thylakoid membrane. The three-dimensional solution structure of this protein from the cyanobacterium *Synechococcus* sp. strain PCC 7002 was determined at pH 5.8 and room temperature using over 900 experimental restraints derived from two- and three-dimensional NMR experiments. The structure is comprised of a well-defined five-stranded β-sheet with (+1, +1, +1, $-4x$) topology. There is no helical region except for a single turn of 3_{10} helix between the βD and βE strands. PsaE also exhibits a large unrestrained loop spanning residues 42–56. A comparison to known protein structures revealed similarity with the Src homology 3 (SH3) domain, a membrane-associated protein involved in signal transduction in eukaryotes. The match is remarkable as 47 of the α-carbons of PsaE can be superimposed onto those of the SH3 domain from chicken brain α-spectrin with a root-mean-square deviation of 2.3 Å. Although the amino acid sequences have low identity and the loops are different in both proteins, the topology of the β-sheet and the 3_{10} turn is conserved. SH3 domains from other sources show a similar structural homology. The structure of PsaE was used to suggest approaches for elucidating its roles within photosystem I.

The photosystem I (PS I)[1] reaction center of higher plants, eukaryotic algae, and the prokaryotic cyanobacteria acts as a light-driven oxidoreductase. PS I utilizes the energy of a single red photon (1.8 eV) to drive the energetically unfavorable reduction of oxidized ferredoxin (E_m = −420 mV) by reduced plastocyanin or cytochrome c_6 (E_m = +370 mV) [for reviews, see Bryant (1992), Chitnis and Nelson (1991), Golbeck (1992, 1994), and Golbeck and Bryant (1991)]. Remarkably, the quantum yield for stable photochemical charge separation across the thylakoid membrane approaches 1.0, and approximately 40% of the energy of the photon is converted into chemical free energy. The reaction center that carries out this energy conversion process is complex and highly conserved. In cyanobacteria, it is made up of about 100 chlorophyll *a* molecules, 10–15 β-carotene molecules, 2 phylloquinone molecules (vitamin K_1), and three [4Fe-4S] clusters denoted F_X, F_A, and F_B. The PS I complex also contains single copies of 11 polypeptides, denoted PsaA, PsaB, PsaC, PsaD, PsaE, PsaF, PsaI, PsaJ, PsaK, PsaL, and PsaM. The genes encoding all of these subunits have been cloned from several cyanobacteria and higher plants (Bryant, 1992; Golbeck, 1992, 1994, and references therein).

The structural and functional roles for many of these protein subunits in PS I have been revealed through a resolution and reconstitution approach (Li et al., 1991; Parrett et al., 1990; Zhao et al., 1992) and through "genetic resolution" of the complex, *i.e.*, through the creation and characterization of interposon mutants in which specific polypeptides are missing from the complex. In particular, the PsaA and PsaB polypeptides form the heterodimeric reaction core and bind most of the chlorophylls and the electron transport components P700, A_0, and A_1 (phylloquinone) and the [4Fe-4S] cluster F_X [see Golbeck (1992) and Golbeck and Bryant (1991)]. The PsaC subunit binds the F_A and F_B [Fe-S] centers of the complex that serve as the terminal electron acceptors (Li et al., 1991; Yu et al., 1993a; Zhao et al., 1992). The PsaD subunit is thought to be involved in the stable binding and proper orientation of PsaC on PS I and the docking of soluble ferredoxin to the complex (Li et al., 1991; Wynn et al., 1989).

Cyanobacterial PsaE proteins contain 69–75 amino acid residues and are strongly basic. With PsaC and PsaD, PsaE forms a protruding cap on the stromal surface of the PS I complex (Böttcher et al., 1992; Kruip et al., 1993). Although mutant strains carrying insertionally inactivated *psaE* genes were originally reported to have no discernible phenotype (Bryant et al., 1990; Chitnis et al., 1989), recent studies revealed that PsaE was required for growth at low light intensity or atmospheric carbon dioxide levels (Zhao et al., 1993). This phenotype suggested that the PsaE subunit might play a role in cyclic electron transport. During cyclic electron transport, electrons from the reducing side of the PS I complex are returned to $P700^+$ via the cytochrome b_6f complex by an

[†] This work was supported in part by National Science Foundation Grant MCB-920576 (D.A.B.) and by National Institutes of Health Grant DK-43101 (J.T.J.L.).

[‡] The coordinates of PsaE have been deposited in the Brookhaven Protein Data Bank under the file name 1PSE.

* Address correspondence to D.A.B. regarding the biochemistry and molecular biology of photosystem I and to C.J.F. or J.T.J.L. regarding NMR matters and structure determination.

[§] Department of Chemistry.

[∥] Department of Biochemistry and Molecular Biology.

[⊗] Abstract published in *Advance ACS Abstracts,* April 15, 1994.

[1] Abbreviations: 1D, one dimensional; 2D, two dimensional; 3D, three dimensional; COSY, correlation spectroscopy; DG, distance geometry; FNR, ferredoxin:$NADP^+$ oxidoreductase; HSQC, heteronuclear single-quantum coherence; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser enhancement spectroscopy; PDB, protein data bank; PS, photosystem; PsaE, photosystem I accessory protein E; 2QF-COSY, double-quantum-filtered correlation spectroscopy; rms, root mean square; rmsd, root-mean-square deviation; SA, simulated annealing; SH2, Src homology 2; SH3, Src homology 3; TOCSY, totally correlated spectroscopy.

0006-2960/94/0433-6052$04.50/0

unknown mechanism. This cyclic flow of electrons allows ATP synthesis to occur from the energy stored in the electrochemical proton gradient that is established. Electron transport studies in the intact cyanobacterial cells confirmed that PsaE is required for cyclic electron transport (Yu et al., 1992b, 1993b).

The nuclear-encoded PsaE proteins of higher plants are significantly larger (91–101 amino acid residues) than their cyanobacterial homologues (Bryant, 1992; Golbeck, 1992, 1994). Cross-linking studies with PS I complexes from higher plants suggest that higher plant PsaE interacts with ferredoxin: $NADP^+$ oxidoreductase (Andersen et al., 1992). Other studies have demonstrated that PsaE may play a direct or indirect role in promoting an interaction between PsaC and its electron-accepting partner, soluble ferredoxin (Rousseau et al., 1993; Sonoike et al., 1993; Weber & Strotmann, 1993).

A major objective of current research on PS I is to understand the structural basis for the remarkably efficient light-energy conversion properties of this enzyme complex. Toward this goal, crystallization of the PS I complex from the thermophilic cyanobacterium *Synechococcus* sp. has been reported (Krauss et al., 1993). However, a high-resolution structural model of the PS I complex may not be available for several years. The ability to produce large amounts of individual subunits allows for the collection of complementary information. Thus, the small PsaE protein was studied in aqueous solution by NMR spectroscopy. The complete assignments of the ^1H and ^{15}N spectra for PsaE (Falzone et al., 1994) provided the necessary background data for the structural determination reported here. A structural homology search showed that PsaE has a fold similar to that of the Src homology 3 (SH3) domain of eukaryotic signaling proteins (Fry et al., 1993; Koch et al., 1991).

MATERIALS AND METHODS

Protein Overproduction and Preparation. PsaE and uniformly ^{15}N-labeled PsaE were prepared and purified as reported elsewhere (Falzone et al., 1994; Zhao et al., 1993).

NMR Spectroscopy. All NMR spectra were collected on 2 mM protein samples at pH 5.8 and 298 K with a modified Bruker AM-500 spectrometer as described in the preceding paper (Falzone et al., 1994). The NMR solution structure calculations were based on 70-ms NOESY spectra in both H_2O and ^2H_2O, a 120-ms 3D ^{15}N NOESY–HSQC experiment, and a 150-ms jump–return NOESY data set. The jump–return experiment provided seven NOEs bleached by presaturation of the intense water resonance in the 70-ms NOESY. A series of J-modulated HSQC experiments yielded backbone ϕ angle restraints, and a 2D H_NNH_{AB} COSY data set was used in the determination of ^{15}N–H_β coupling constants for stereospecific assignments (Chary et al., 1991). Details of the acquisition and processing parameters for the above experiments are given in Falzone et al. (1994).

Interproton Distance Restraints. A total of 853 interproton restraints from NOESY experiments were used to calculate three-dimensional solution structures. Of these, 378 were intraresidue and 475 were interresidue. A total of 230 restraints were long range ($|i-j| > 5$), 71 were short range ($1 < |i-j| \leq 5$), and 174 were sequential ($|i-j| = 1$), where i and j represent amino acid numbers. Approximate interproton distances were calibrated with the $C^\delta H_2$–$C^\epsilon H_2$ Tyr cross-peak volume (divided by two) or the single Trp $C^{\zeta 2}H$–$C^\eta H$ cross-peak volume. The observed NOEs were assigned to three categories corresponding to strong, medium, or weak intensities. The respective distance ranges for these three categories were 1.8–2.7, 1.8–3.3, and 1.8–5.5 Å. Methyl groups were treated as X-PLOR pseudoatoms (Brünger, 1992) with $\langle r^{-6} \rangle$ averaging; 0.5 Å was added to the upper bound (Wagner et al., 1987) and the allowable lower limit was unchanged. Likewise, no distinction was made between $C^{\delta 1}H$ and $C^{\delta 2}H$ ($C^{\epsilon 1}H$ and $C^{\epsilon 2}H$) of Tyr or Phe rings, and $\langle r^{-6} \rangle$ averaging was applied.

Dihedral Angle Restraints. Backbone ϕ values were obtained as described in Falzone et al. (1994) and were based on a series of J-modulated HSQC experiments. A total of 34 ϕ angles were placed in two categories: $-50 \pm 40°$ when $^3J_{\alpha H\text{-}H} < 6$ Hz and $-130 \pm 50°$ when $^3J_{\alpha H\text{-}NH} > 8$ Hz. However, a single ^3J generally corresponds to multiple angle values. During the refinement, three angles giving rise to systematic violations were moved from their original category: Ser-26, from $-50°$ to $-150°$; Asn-59 and Ala-61, from $-130°$ to $-90°$. A total of 29 χ_1 angle constraints were determined on the basis of 2D H_NNH_{AB} COSY data, 3D ^{15}N NOESY–HSQC data, 3D ^{15}N TOCSY–HSQC data, and 2QF-COSY data (Clore & Gronenborn, 1989): 22 were assigned to $-60 \pm 30°$, 4 to $180 \pm 30°$, and 3 to $60 \pm 30°$.

Hydrogen-Bonding Restraints. The presence of hydrogen bonds was deduced from a series of ^1H–^{15}N HSQC spectra of freshly dissolved PsaE in ^2H_2O (Falzone et al., 1994). Cross peaks that remain in these spectra are indicative of amide protons protected from exchange with the bulk solvent either by hydrogen bonding or by sequestration of the proton within the protein structure. A total of 18 cross peaks were found in the ^1H–^{15}N HSQC spectrum after about 5 h (experiment collected from 4 to 5 h after dissolution of PsaE in ^2H_2O) in ^2H_2O whereas 9 were present after 8 h. Of the 18 observed amide resonances, 17 were used as constraint inputs in the structure calculations. The structures generated just prior to adding these constraints indicated the hydrogen-bonding partners through amide-to-carbonyl distances of ca. 2 Å. The amide proton in these hydrogen bonds was set to 1.5–2.3 Å from the acceptor carbonyl oxygen, and the amide nitrogen to acceptor atom distance was restrained within 2.5–3.3 Å.

Iterative Structure Refinement. An iterative strategy was applied as suggested in Powers et al. (1993). Initial structures were obtained by using a small set of clearly assigned NOEs; other restraints were confirmed and added in successive rounds of calculations. Backbone dihedral constraints (ϕ) were included after about 70% of the assigned NOEs had been used; side-chain χ_1 constraints, after about 90%. Hydrogen-bonding restraints were taken into consideration in the final rounds of calculation, which included all NOE and dihedral constraints. With the assignment of three interproton distance categories, some NOEs fall near the cutoffs. These restraints produce systematic distance violations stemming from their assignment to the more restrictive distance cutoff and were placed in the next interproton distance category (Powers et al., 1993), a procedure that does not compromise the quality of the structure (Clore et al., 1993).

Structural Calculations. All three-dimensional structures were calculated by using the program X-PLOR 3.1 (Brünger, 1992) with the protocol based on the hybrid distance geometry–simulated annealing (DG–SA) method of Nilges et al. (1988). The target energy function in both restrained conjugate gradient minimization and dynamical simulated annealing consists of energy terms to maintain the covalent geometry (*i.e.*, bond lengths, angles, planes, and chirality), square-well potential terms for the NOE and dihedral angle restraints, and a repel potential term for the van der Waals nonbonded contacts. No other noncovalent forces such as electrostatic

interaction, hydrogen bonding, or 6–12 Lennard-Jones empirical potential were included in the target function. A total of 100 initial substructures of PsaE were calculated by embedding the bound distances for 43% of the atoms in the Cartesian coordinate space, that is, by transforming the interatomic distances into Cartesian coordinates (Havel & Wüthrich, 1984; Havel et al., 1983); this was followed by 100 steps of Powell minimization against *only* the effective energy term for distance bounds. The missing atoms were then added by fitting the ideal side-chain geometry at each residue. After 200 steps of restrained minimization, a dynamical simulated annealing was carried out on these full structures: 1000 steps (Δt = 0.003 ps) of Verlet dynamics at 2000 K with slow reduction of van der Waals repulsion and increase of chirality and planarity energy terms, followed by 1000 slow cooling steps (Δt = 0.005 ps; ΔT = 50 K per 26 steps) to the final temperature of 100 K with gradual increase of the van der Waals energy terms. To ensure the consistency of the final structures and to compare the efficiency of different algorithms, three other methods were also applied: full structure distance geometry with partial random metrization (Kuszewski et al., 1992), *ab initio* simulated annealing (Nilges et al., 1991), and random simulated annealing (Nilges et al., 1988). All methods yielded equivalent results; however, in terms of the computation time and the number of accepted structures, DG–SA and *ab initio* SA were more efficient than the other algorithms in the case of PsaE.

All structures were subjected to an SA refinement that consisted of 2000 slow cooling steps (Δt = 0.005 ps; ΔT = 50 K per 111 steps) from the initial annealing temperature of 1000 K to the final temperature of 100 K. This was followed by 300 steps of restrained minimization with effective van der Waals radii set to 0.75 times the standard values used in the CHARMm empirical energy function (Brooks et al., 1983). Subsequent SA refinements and minimizations used a factor of 0.8. The majority of the resulting structures showed no NOE violations greater than 0.2 Å and no dihedral constraint violations greater than 3°. On the basis of the total energy and the number of violations (<2), 76 conformations remained in the iterative refinement procedure, consisting of SA and 2000–5000 steps of the restrained minimization. The extensive minimization restored some slightly strained bond angles without violating the experimental constraints and without destroying the adequate conformational sampling obtained in the DG–SA algorithm. Of the 76 structures, 30 were found to have no NOE violations larger than 0.1 Å and no dihedral violation larger than 1°.

RESULTS AND DISCUSSION

Analysis of the NOEs, backbone *J*-coupling constants, and amide H-exchange rates revealed that the PsaE protein contains an antiparallel five-stranded β-sheet, connecting turns, and no helical regions (Falzone et al., 1994). The spectral information was used to produce a set of constraints and to construct a three-dimensional model of the protein in solution.

Final Structures. The final structures were derived from the DG–SA method. The final round of energy minimization returned a set of 30 structures compatible with all constraints within 0.1 Å and 1°. These structures contain no bad contacts as indicated by low effective van der Waals energies and negative Lennard-Jones energies calculated with CHARMm parameters (Brooks et al., 1983). The 30 structures were aligned with a least-squares fit on the residues forming the regular secondary structure (see below), and from the aligned set, a mathematical average structure was calculated. Each member of the family was compared against this average, and the corresponding statistics are listed in Table 1. Figure 1A illustrates the N–C^{α}–C–O traces of these 30 aligned structures.

Table 1: Structural Statistics and Atomic rms Deviations

Structural Statistics	{SA}[a]	⟨SA⟩$_r$[b]
rms deviation from experimental constraints		
distance constraints (Å) (887)[c]	0.0086 ± 0.0009	0.0095
dihedral constraints (deg) (63)[d]	0.11 ± 0.02	0.13
rms deviation from ideal geometry		
bonds (Å)	0.0028 ± 0.0001	0.0027
angles (deg)	0.42 ± 0.01	0.41
impropers (deg)	0.242 ± 0.008	0.238
no. of violations[e]		
distance constraints	78 ± 4	75
dihedral constraints	6 ± 1	7

rms Deviations from the Average Structure (Å)[f]	all residues	limited set[g]	regular 2° structure[h]
backbone (N–C^{α}–C)	1.73 ± 0.54	0.34 ± 0.07	0.19 ± 0.04
side chains (heavy atom)	1.77 ± 0.34	0.91 ± 0.12	0.76 ± 0.06

[a] {SA} is the family of the final 30 simulated annealing structures; mean values and standard deviations are shown for these 30 structures. [b] ⟨SA⟩$_r$ is the structure obtained by applying 5000 steps of restrained minimization to the mathematical average structure. The mathematical average structure was obtained by first aligning the 30 structures with a least-squares fit over the regular secondary structure [(N–C^{α}–C) of residues 6–10, 21–26, 36–39, and 57–68] and then individually averaging the atomic coordinates. [c] A total of 887 distance constraints were used in the structure calculation: 230 were long range ($|i-j| > 5$), 71 were short range ($1 < |i-j| \leq 5$), 174 were sequential ($|i-j| = 1$), 378 were intraresidue, and 17 × 2 were H-bond constraints. [d] A total of 63 dihedral constraints were used in the structure calculation: 34 were backbone φ angles, and 29 were side-chain χ_1 angles. [e] *All* distance violations are less than 0.1 Å, and *all* dihedral violations are less than 1°. [f] The rms deviations of the atomic positions were calculated for each structure in {SA} against the mathematical average structure. The average values and standard deviations for the 30 final structures are reported. [g] rmsd for residues 2–27, 34–40, and 57–69. This set excludes the disordered BC and CD loops and the N-terminal residue. [h] rmsd of the segment adopting regular secondary structure: residues 6–10, 21–26, 36–39, and 57–68.

Figure 2 presents a histogram of the rms deviation from the average structure versus residue number for backbone atoms (N–C^{α}–C, solid columns). Overall, the backbone atoms exhibit an rms deviation of 1.73 Å from the average structure; however, Figure 2 shows clearly that the deviation is not uniform over the sequence. The backbone atoms within ranges 2–27, 34–40, and 57–69, that is, portions of the protein not including two disordered loops, have an rms deviation of 0.34 Å. The subensemble comprising the segments that adopt regular secondary structure (6–10, 21–26, 36–39, and 57–68) has an average rms deviation of 0.19 Å, a value reflecting the high density of restraints in these regions of PsaE (>15 restraints per residue for amino acids spanning 1–42 and 57–69). Outside the well-restrained regions, the backbone atoms exhibit an average rms deviation larger than 1 Å, reaching 6 Å at the tip of the loop extending from Lys-42 to Asn-56.

Side-chain rms deviations for all heavy atoms are represented with open columns in Figure 2. The value for the entire molecule is 1.77 Å, but when the two disordered loops are excluded, this value drops to 0.91 Å. As expected, side chains found at the surface display a greater deviation from the average structure and are the main contributors to the 0.91-Å value. In general, side-chain atoms follow the trend of the backbone atoms: the residues from the ill-defined loops are found in an ill-defined conformation, in sharp contrast to the buried side chains appended to those regions of the backbone corresponding to the β-sheet. For several long side chains, *e.g.*, Arg-4 and Arg-39, there is a discrepancy between the

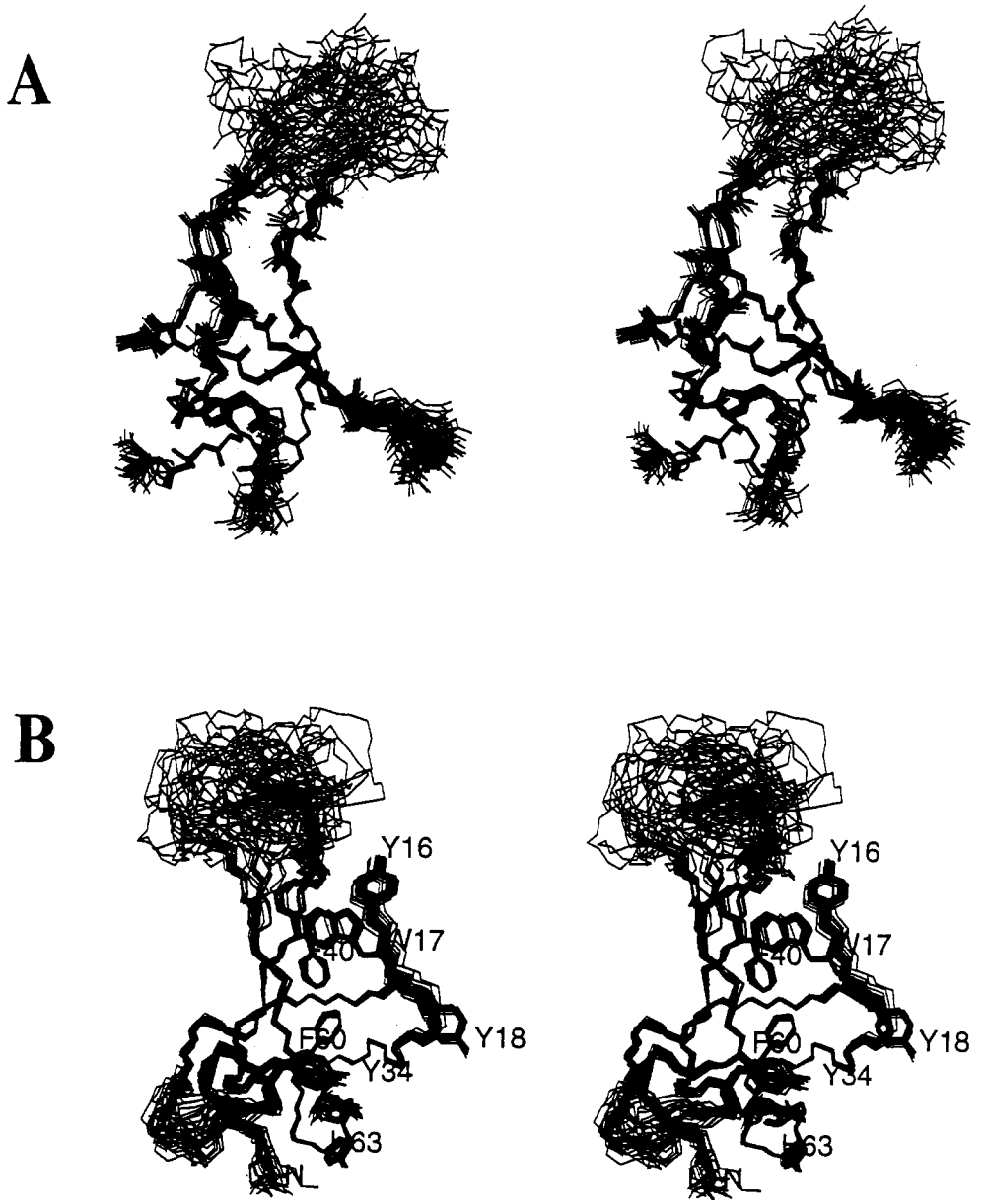

FIGURE 1: (A) Stereoview of the best fit superposition of 30 PsaE structures calculated using 950 experimentally derived restraints. The fit was performed on the well-defined regions of the structure as described in Table 1, and only the backbone is represented. The C-terminus is at the bottom, on the left. The PsaE protein consists of a five-stranded β-sheet and a turn of 3_{10} helix between βD and βE. The backbone atoms are highly restrained in these regions of the protein. Three loops connect the first four strands: the AB loop is well-defined by experimental restraints as it contains three aromatic residues that make long-range contacts. The BC and the CD loops (27–34 and 42–56, respectively) are less well-defined. The CD loop, which has few long-range NOEs, displays large rms deviations from the average structure. (B) Stereoview of the same 30 structures with the aromatic side chains included to show part of the hydrophobic interior and the range of deviations from the average structure for these residues. The N-terminus is at the bottom, in the middle.

backbone and the side-chain rmsd values. Inspection of the family of structures reveals that a 2-fold degeneracy of χ dihedral angles, resulting from the lack of proper constraint beyond the β-carbon, is responsible for the peculiarity.

The backbone displays allowed (ϕ, ψ) values, except for glycines and occasionally some of the residues in poorly constrained segments. Only one site (Asn-41) appears abnormal as it exhibits a mixture of positive and negative ϕ angles in the final structures. This occurs because the C$^{\alpha}$H of 41 resonates at 4.74 ppm, overlapping with the C$^{\alpha}$H of Phe-40 and the intense H_2O signal. Although a weak cross peak is present in the jump–return NOESY at the amide chemical shift of Asn-41 and 4.74 ppm, it is not possible to determine its true intensity and its origin as the αN(i, i+1) NOE for Phe-40 and Asn-41. The gap in sequential backbone NOEs at the start of the ill-defined 42–56 region allows both confomers to be populated; although the low intensity of the sequential cross peak is consistent with a negative ϕ value, the ambiguity will remain until nondegenerate data can be obtained and will contribute to the rmsd in this region.

Structure of PsaE. (A) Regular Structure. Figure 3 shows a ribbon structure obtained from one member of the family displayed in Figure 1. The high-rmsd loop conformations presented in Figure 3 are consistent with the NMR constraints, but as can be seen in Figure 1, they are not uniquely fixed. The main structural motif of the protein is the β-sheet (Falzone et al., 1994). The βA strand (residues 7–10) is hydrogen bonded to both the βB strand (residues 21–26) and the βE strand (residues 66–68). The β-sheet hydrogen-bonding pattern between βA and βB is broken by a bulge at Gly-22; cross-strand NOEs are restored with the C$^{\alpha}$H of Val-21 to the NH of Val-8 and the C$^{\alpha}$H of Lys-9. In addition, the structure

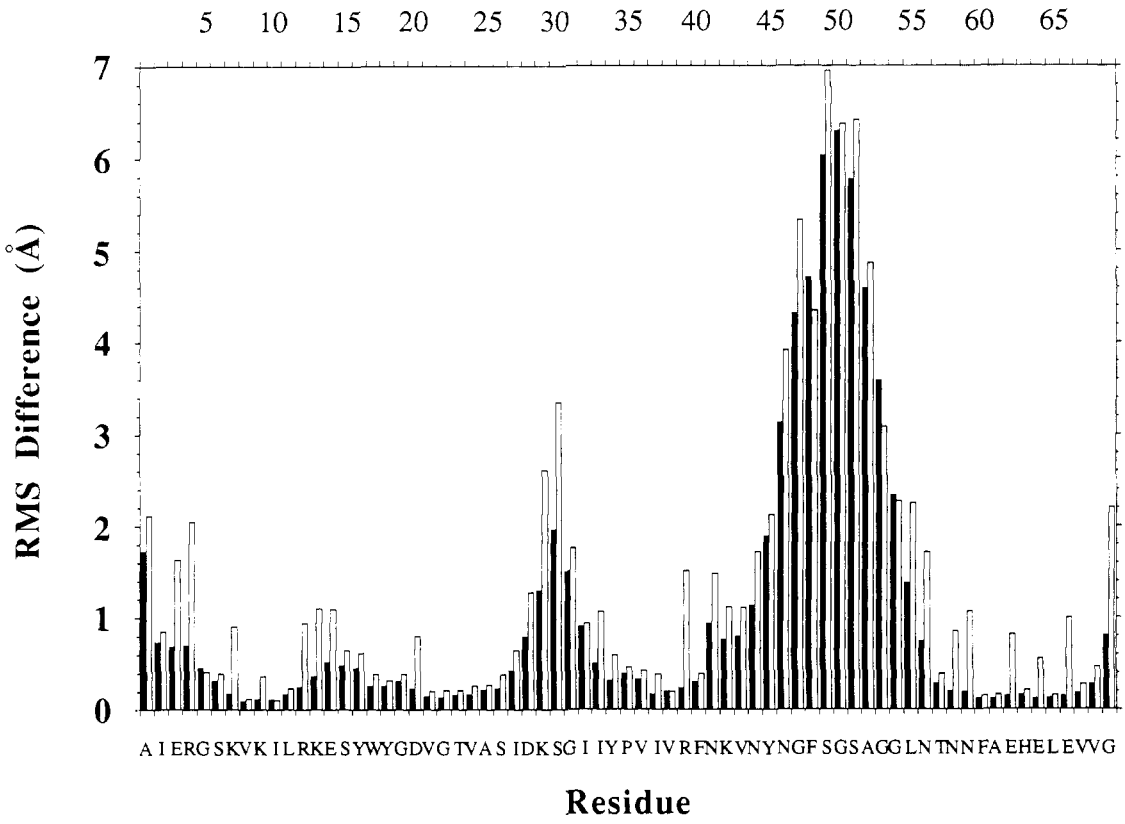

FIGURE 2: Root-mean-square deviations from the mathematical average structure. Solid bars represent the backbone atoms (N–C$^{\alpha}$–C), and open bars represent all heavy atoms within each residue. The well-defined β-strands (7–10, 21–26, 36–39, 57–59, and 66–68) and the 3_{10} turn (62–65) are readily recognized by the low rms deviation of backbone and nonsurface side-chain atoms. The large CD loop shows large rms deviations for both backbone and side-chain atoms as it is underconstrained. The smaller BC loop, which also has few long-range restraints, has rmsd values typical of a loop of its size.

confirms that Asp-20 provides the accepting carbonyl in a hydrogen bond involving the NH of Ile-10. The βC strand (residues 36–39) and the βD strand (residues 58–60) complete the β-sheet. The proximity of βA and βE stems from the right-handed twist of the β-sheet.

Little is known about the importance of β-sheet integrity for the folding and function of PsaE, but it has been observed that introduction of a stop codon at position 65, eliminating the last strand, does not allow binding of PsaE to PS I (Rousseau et al., 1993). Several cyanobacterial proteins have C-terminal extensions that are enriched in alanine, proline, and lysine (Bryant, 1992). In contrast, the longer PsaE subunits of higher plants have sequences at the C-terminal end of their N-terminal extension (near the beginning of the conserved sequence domain) that are enriched in the same amino acids. The proximity of the two termini in *Synechococcus* sp. strain PCC 7002 PsaE suggests that, in the longer cyanobacterial proteins and in the higher plant proteins, it is likely that the additional positive charges are found in approximately the same position with respect to the rest of the PS I complex.

(B) Loops. The three loops connecting the strands of the β-sheet are distinct in their appearance. The first connects βA with βB, stretching from Leu-11 to Gly-19; it satisfies the distance requirements of Ω-loops (Leszczynski & Rose, 1986). The AB loop contains an aromatic triplet (Tyr-16, Trp-17, Tyr-18) illustrated in Figure 1B. These three residues form the boundary of a cluster on which nonpolar side chains dock and give rise to many long-range NOEs. As a consequence of this clustering, the AB loop assumes a relatively well-ordered structure. While residues 13–17 in the AB loop are strictly conserved (Zhao et al., 1993), the tyrosine at position 18 is occasionally replaced by a phenylalanine. Gly-19 can also be substituted with Gln, Asn, or Lys. Rousseau et al. (1993) have made several site-specifically altered PsaE proteins in *Synechocystis* sp. strain PCC 6803. Changing Trp-17 to Ala, or changing Tyr-16 and Tyr-18 both to Phe, did not prevent folding and binding of the protein to PS I.

The second loop (BC loop) connects βB and βC and extends from Ile-27 to Pro-35. As illustrated in Figures 1 and 2, the BC loop is more disordered than the AB loop and is protruding into solution. The lysine at position 29 is found only in *Synechococcus* sp. strain PCC 7002 and *Synechocystis* sp. strain PCC 6803. In these two PsaE's, position 33 is occupied by an aliphatic residue (isoleucine or leucine). In contrast, in most known PsaE variants, position 29 is occupied by an uncharged residue but position 33 by a basic residue. Rousseau et al. (1993) have observed that in *Synechocystis* sp. strain PCC 6803 Lys-29 is the only residue available for reaction when the whole PS I complex is subjected to fluorescein isothiocyanate. This observation suggests that the BC loop of cyanobacterial PsaE is accessible to solvent and that the basic residue has a unique behavior which might have some functional importance.

The third and last loop is located between βC and βD. This loop is perhaps the most striking feature of the structure; it is large relative to the size of the protein, extending over 14 residues. The lack of long-range NOEs results in underconstraining and disordered appearance. Preliminary ^{15}N relaxation studies (C. J. Falzone, Y.-H. Kao, and J. T. J. Lecomte, unpublished results), to be part of a full dynamic analysis, point to longer longitudinal relaxation times in this loop than on the average for the whole protein and suggest that this part of the structure is more mobile than the rest of the backbone. Increased mobility is also likely at the

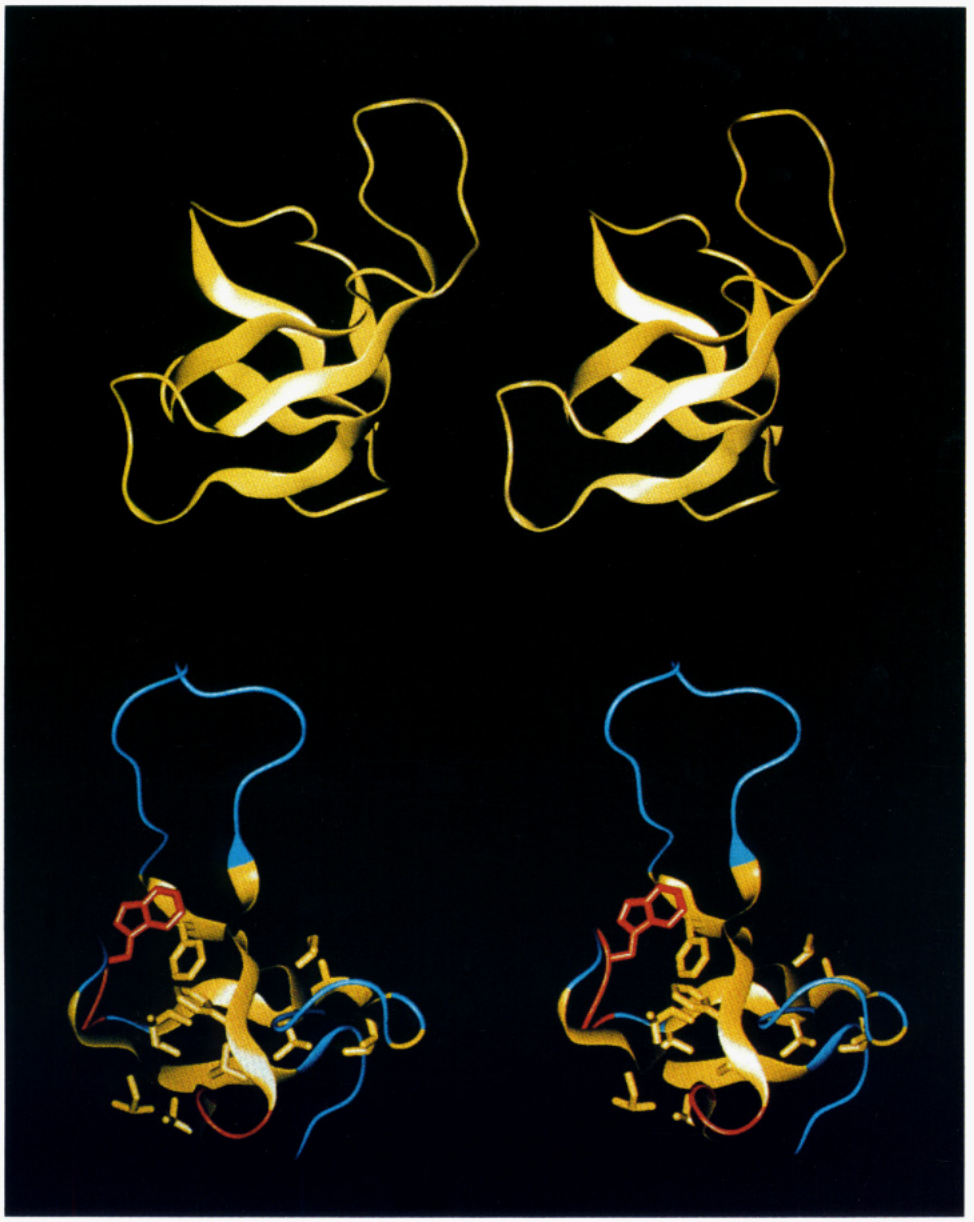

FIGURE 3: Stereo ribbon representation of one of the 30 PsaE structures shown in Figure 1. With this topology and orientation, the N- and C-terminus are brought together at the bottom of the panel. Top pair: The N-terminus is at the bottom, in the center. Strand βA starts at the bottom, on the right, and lies at 45° in the back of the structure. The bulge at the beginning of strand βB is visible at the right edge. Strand βC leads into the CD loop, which points up. Strand βD is the top strand, in the front of the structure, and ends in the 3_{10} turn, on the left. Strand βE is horizontal in the back. The β-sheet scheme is provided in Falzone et al. (1994). Bottom pair: A different orientation of the ribbon structure displaying several of the constrained hydrophobic residues: Val-8, Ile-10, Leu-11, Trp-17, Val-21, Val-24, Ala-25, Ile-27, Ile-37, Phe-40, Phe-60, and Val-68. The barrel-like five-stranded (+1, +1, +1, $-4x$) β-sheet is in yellow; the 3_{10} turn (bottom, front) and AB loop (left, front) are in brown.

N-terminus where longer relaxation times are observed as well.

The primary structure of PsaE proteins from several other species lacks seven residues corresponding to Gly-50 to Asn-56 in *Synechococcus* sp. strain PCC 7002 PsaE. These species include *Nostoc* sp. strain PCC 8009, *Calothrix* sp. strain PCC 7601, *Porphyra umbilicalis*, spinach, barley, and *Chlamydonomas reinhardtii* (Zhao et al., 1993). However, it should be noted that Tyr-45 and Phe-48 within the CD loop are not deleted in any known primary structure. In fact, Tyr-45 is completely conserved while Phe-48 is occasionally replaced by a tyrosine, valine, or isoleucine. The role of these amino acids is yet to be understood.

The connection between βD and βE is made by a relatively small traverse taken by residues 61 (still in extended conformation) through 65 before the final β-strand. Residues 62–65 form a type III turn (3_{10} helical turn) and show a stable hydrogen bond between the amide proton of Leu-65 and the carbonyl oxygen of Glu-62. The residues of this turn consist of Glu-62 (may be Asp or Leu in other PsaE proteins), His-63, Glu-64 (conserved), and Leu-65 (may be Val or Ile). Aside from the sequential glutamic acids, the environment of His-63 comprises Ala-61, Ile-33, and Tyr-34. These residues lower the p$K_{1/2}$ of the imidazole group to 5.35 (Falzone et al., 1994).

(C) Distribution of Residues. The various types of residues are not evenly distributed through the structure. Figure 1B displays the aromatic side chains. As observed above, Tyr-45 and Phe-48 are not well constrained and their orientation relative to each other or to the rest of the protein cannot be determined. Tyr-34, at the end of the BC loop, points toward the solution and is best described with several orientations. Tyr-16 and Tyr-18 are largely exposed to solvent and display

more ring disorder than the aromatic rings confined to the core of the protein. Trp-17 is maintained almost parallel to both Tyr-16 and Phe-40; Phe-60 is perpendicular to this line of three aromatics, directly across from Phe-40. Residues 17, 40, and 60, particularly Phe-40, form the heart of the hydrophobic core of PsaE.

PsaE contains a total of 21 aliphatic residues, which are found mostly in the sheet (the bottom stereo pair of Figure 3 displays some of these). Val-43, Ala-52, and Leu-55 are the only ones found in the CD loop; they are not engaged in strong interactions with each other or with other residues. Other aliphatic residues, which belong to the βA, βC, and βE strands, cluster near the C- and N-termini. There are numerous aromatic–aliphatic contacts in the core; for example, Phe-40 is surrounded by Val-8, Ile-10, and Val-38 (Falzone et al., 1994). As is observed for the interior aromatics, the aliphatic side chains in the hydrophobic core are generally ordered: Val-36, Val-38, and Leu-65 are part of the hydrophobic cluster that includes Phe-40 and Phe-60 and have their side chains well-defined. Other aliphatic side chains also display ordered conformations except those in the BC loop (Ile-32 and Ile-33) and those at the N-terminus (Ala-1 and Ile-2).

Seven acidic residues and eight basic residues are found in PsaE from *Synechococcus* sp. strain PCC 7002. As expected, all of the charged residues are at the surface of the protein where they contribute to the high solubility. No charged residue is located in the large CD loop or along two hydrophobic ridges at the surface of the protein. These two lines of hydrophobic side chains are made up of Ala-1, Ile-2, Val-67, and Val-21 in one set and Ala-25, Ile-37, Ile-32, Ile-33, and Tyr-34 in the other. Pro-35 is just below the surface at the end of this ridge. Ile-27, which is positioned between Ile-2 and Ile-37, bridges these two hydrophobic areas. The bottom part of Figure 3 illustrates this arrangement of hydrophobic side chains in the core of the protein.

Structural Homology with Other Proteins. The coordinates for one of the structures shown in Figure 1 were compared with available three-dimensional structures by using the programs WHAT IF (Vriend, 1990) and DALI (Holm & Sander, 1993). The programs identified strong homology with SH3 domains, in particular that found in α-spectrin from chicken brain (PDB entry 1shg; Musacchio et al., 1992; Noble et al., 1993). The topology of the β-sheet and the 3_{10} turn is conserved, and the best alignment in these regions (47 α-carbons) yields an average α-carbon rms deviation of 2.3 Å. This remarkable match is represented in Figure 4 where the α-carbon traces of both proteins are superimposed. The alignment of the primary structures of PsaE and α-spectrin is described in Table 2. Amino acid sequence identity is only 13%, a value in the twilight zone (Doolittle, 1990; Sander & Schneider, 1991). The α-carbon rms deviations for other SH3 domains available in the protein data bank are 2.6 Å for the SH3 domain of biotin operon repressor (PDB entry 1bia; Wilson et al., 1992), 2.7 Å for phospholipase C_γ (PDB entry 1hsp; Kohda et al., 1993), and 2.8 Å for the p85α subunit of phosphatidylinositol 3-kinase (PDB entry 1pnj; Booker et al., 1993). Table 2 lists the alignment with the SH3 domain from chicken c-Src (Yu et al., 1992a) as reported by Noble et al. (1993). Plasmid-encoded dihydrofolate reductase (R67, coordinates not available) is also known to have the same five-stranded β-sheet and 3_{10} turn as PsaE with an active site formed at the dimer interface of two subunits (Matthews et al., 1986).

SH3 domains, along with SH2 domains, are involved in eukaryotic signal transduction pathways. One of the functions suggested for SH3 domains involves protein–protein interactions at membrane surfaces and at the cytoskeleton (Koch et al., 1991). SH2 and SH3 domains are commonly found within larger proteins containing a catalytic domain such as a kinase or a phospholipase (Koch et al., 1991). Many of these proteins are involved in regulation of cellular proliferation and in differentiation, processes that are mediated by membrane-bound growth factor receptors with inducible protein–tyrosine kinase activity. Phosphorylation of the receptor creates a high-affinity binding site for the SH2 domains. Intracellular proteins, which bind to these receptors, include the tyrosine kinases of Src, Fyn (a member of the Src family), Abl, Fps, and phospholipase Cγ and $p21^{ras}$ GTPase-activating protein (GAP). All of these proteins contain not only SH2 domain(s) but also SH3 domain(s) that may be acting in concert. However, SH3-like domains are often found as independent structural units (Noble et al., 1993), *e.g.*, in α-spectrin and the *Escherichia coli* biotin operon repressor mentioned above. PsaE provides an additional example of an independent SH3 domain.

SH3 domains are known to interact with other proteins. The best described example is that of the SH3 domain from the tyrosine kinase Src, which binds a protein (SH3 binding protein 1 or 3BP-1) through a proline-rich region (Cicchetti et al., 1992), specifically PPPLPP (Ren et al., 1993). Yu et al. (1992a) used NMR spectroscopy to identify the residues of the Src SH3 domain that interact with this sequence. They analyzed the chemical shift perturbations induced by two related peptides and proposed that the contact points (boldface residues in Table 2) involve many of the conserved amino acids in SH3 domains. Alignment of the primary structures of human Fyn and chicken α-spectrin with the Src protein (Noble et al., 1993) allows the corresponding residues to be identified in PsaE, if they exist. They are highlighted in red on the α-spectrin backbone in the bottom stereo pair of Figure 4.

Table 2 illustrates that a few hydrophobic residues in PsaE occupy comparable locations within the proposed SH3 binding site. For example, the sequence VASIDK (24–29) corresponds to LLNSTN in α-spectrin and IVNNTE in the SH3 domain of Src. The PVIV (35–38) stretch in PsaE takes the place of the characteristic WWLA (WWKV in spectrin) segment in SH3 domains. It is possible that Ala-25 and Ile-37, which are part of a hydrophobic ridge in PsaE, participate in accessory protein binding. However, the sequence identity with known SH3 domains is consistently low, the highest values being 16% for the p85α subunit of phosphatidylinositol 3-kinase, and functional relatedness cannot be inferred with any of the structural homologues.

In plants one instance of identical β-sheet topology was found by the superposition programs: actinidin is a two-domain sulfuryl proteinase extracted from the Chinese gooseberry (PDB entry 2act; Baker, 1980). The α-carbon of the C-terminal domain, whose function is so far unexplained, shows a 2.8-Å rmsd match with PsaE. The N-terminal fragment of PsaE located before strand βA is replaced by a helix in actinidin, and the 3_{10} turn, which is strikingly conserved in the other SH3 domains, is extended into a larger loop. The α-carbon trace of PsaE shows homology with other membrane-associated proteins. These include the B-subunits of verotoxin-1 (PDB entry 1bov; Stein et al., 1992) and enterotoxin (PDB entry 1lts; Sixma et al., 1991, 1993). Although these proteins are examples of the oligomer binding (OB) fold

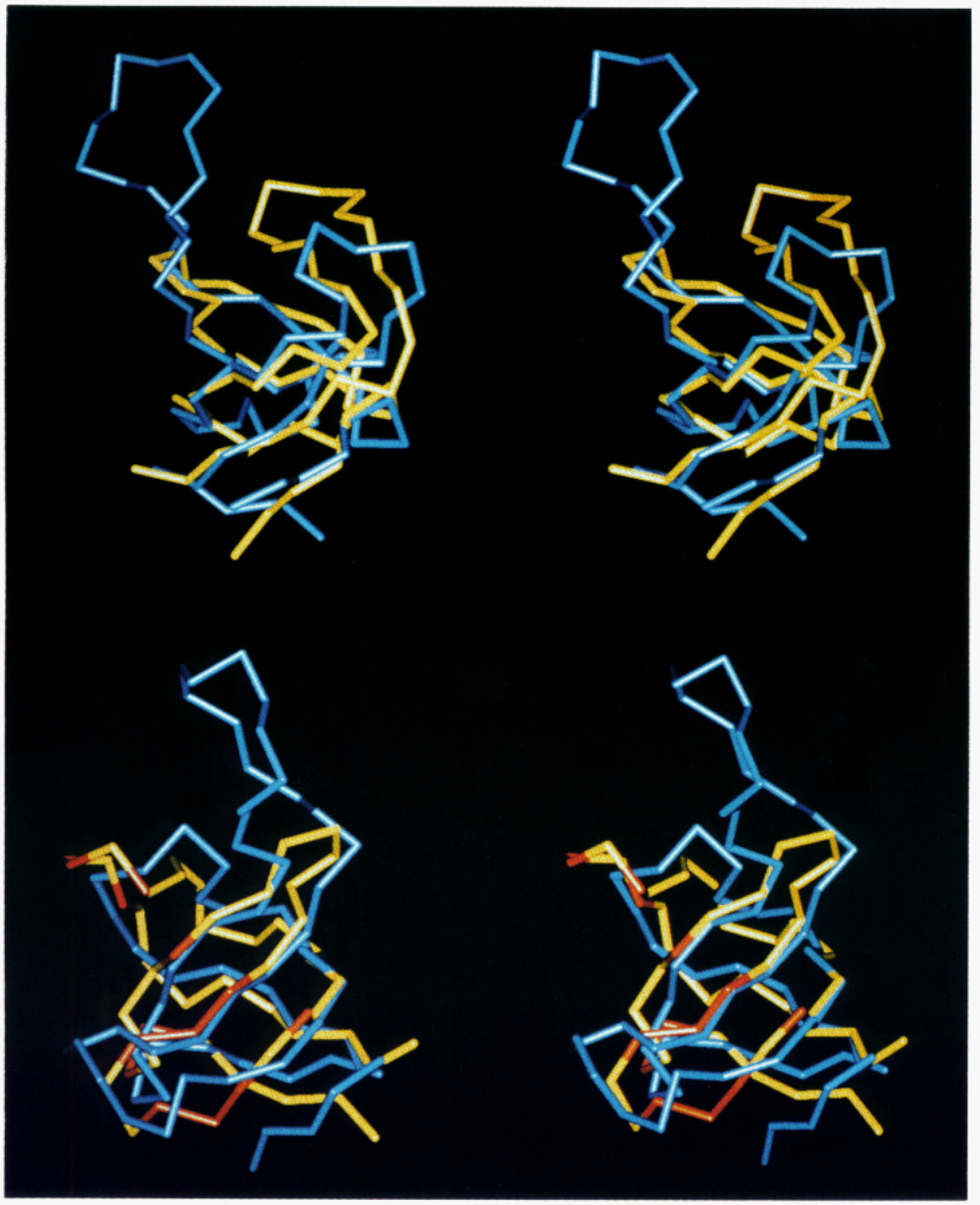

FIGURE 4: Stereoview of the superposition of one of the PsaE structures (in blue) with the X-ray crystal structure of chicken brain α-spectrin (in yellow). These structures were aligned with a least-squares procedure according to the homology data given in Table 2. The α-carbons of the five strands of the β-sheet and 3_{10} turn show a 2.3-Å rms deviation. PsaE and spectrin have nonmatching residues in the connecting loops. The bottom pair presents a different viewing angle, with the ligand binding site proposed for the Src SH3 domain (Yu et al., 1992a) identified in red (boldface residues in Table 2).

(Murzin & Chothia, 1992) and have a permuted topology that renders the significance of the match difficult to evaluate, it is interesting that the spatial arrangement of the backbone is somewhat conserved in these membrane-associated proteins. The evolutionary implications of these intriguing structural homologies will be better understood once the functions of the SH3 domains and PsaE are defined.

Roles of PsaE in Photosystem I Structure and Function. Three roles have been suggested for PsaE in PS I structure and function: (1) binding of ferredoxin:$NADP^+$ oxidoreductase (FNR) to PS I in higher plants; (2) stimulation of the reduction of soluble electron acceptors such as ferredoxin or flavodoxin or both; (3) an unspecified but required role in cyclic electron transport around PS I. The availability of a three-dimensional structure prompts the reinspection of these roles and suggests new directions in the study of the PsaE subunit and PS I.

The PsaE proteins from the higher plants spinach and barley are considerably larger (91–101 amino acid residues, respectively) than the cyanobacterial proteins [69–75 amino acids; for sequence alignments, see Bryant (1992) and Golbeck (1994)]. Andersen et al. (1992) have reported the isolation of a PS I preparation from barley that contains about 0.4 molar equiv of FNR relative to P700. The FNR in such preparations could be cross-linked to the PsaE subunit. In spinach PS I preparations that contain no FNR, the N-terminal region of PsaE (up to residue 15) is sensitive to proteases (Lagoutte & Vallon, 1992). This observation contrasts with that for PsaE in the cyanobacterium *Synechocystis* sp. strain PCC 6803, for which both the N- and C-termini were shown to be buried in PS I preparations (Rousseau et al., 1993). Therefore, it is possible to speculate that in higher plants the N-terminal domain of PsaE is providing a specific site of interaction with FNR. It should be noted that in the cyanobacterium *Synechococcus* sp. strain PCC 7002 FNR is not associated with PS I (U. Mühlenhoff and D. A. Bryant, unpublished results) but is associated with the peripheral rods of phycobilisomes by means of the CpcD domain that occurs

Table 2: Sequences of PsaE, Chicken Brain α-Spectrin, and Chicken c-Src SH3 Domains

	1				5					10					15									
PsaE[a]		A	I	E	R	G	S	K	V	K	I	L	R	K	E	S	–	–	–	–	–	–	–	Y	W
					t	t		β	β	β	β			t	t									t	t
Score[b]						4	7	7	6	6	4	4	3	4	0	1								2	3
α-Spectrin[c]	M	D	L	T	G	K	E	L	V	L	A	L	Y	D	Y	Q	E	K	S	P	R	E	V	T	M
								β	β	β	β														
Src SH3[d]								T	F	V	A	L	**Y**	D	Y	E	S	**R**	**T**	E	**T**	D	L	S	F
								β	β	β	β	β	β	β	β							β	β	β	β

				20					25				30					35					40	
PsaE	Y	G	–	D	V	G	T	V	A	S	I	D	K	S	G	I	I	Y	P	V	I	V	R	F	N
	t	t			β	β	β	β	β	β	e									β	β	β	β		
Score	4	2		5	6	6	7	6	6	6	4	3	1	4					4	5	6	6	5	6	6
α-Spectrin	K	K	G	D	I	L	T	L	L	N	S	T	N	K	D	–	–	–	W	W	K	V	E	V	N
					β	β	β	β	β	β									β	β	β	β	β	β	
Src SH3	K	K	G	E	R	L	Q	I	**V**	N	**N**	**T**	**E**	G	D	–	–	–	**W**	**W**	**L**	A	H	S	L
					β	β	β	β	β										β	β	β	β	β	β	β

						45					50					55					60				
PsaE	–	–	K	V	N	Y	N	G	F	S	G	S	A	G	G	L	N	T	N	N	F	A	E	H	E
																		β	β	β	β	e	3_{10}	3_{10}	3_{10}
Score			3														6	5	3	5	4	5	4	5	5
α-Spectrin	–	–	D	–	–	–	–	–	–	–	–	–	–	–	–	–	R	Q	G	F	V	P	A	A	Y
																	β	β	β	β	β		3_{10}	3_{10}	3_{10}
Src SH3	T	T	G	–	–	–	–	–	–	–	–	–	–	–	–	–	Q	T	G	**Y**	I	P	S	**N**	**Y**
																	β	β	β	β	β		3_{10}	3_{10}	3_{10}

	65				
PsaE	L	E	V	V	G
	β	β	β	e	
Score	5	6	6	5	
α-Spectrin	V	K	K	L	D
	β	β	β	β	
Src SH3	V	A	P	S	
	β	β	β	β	

[a] The primary structure of PsaE with the residue number above and the secondary structure assignments below, according to the H-exchange protection data (Falzone et al., 1994) and dihedral angles. Symbols are as follows: β, participation in the β-sheet; e, extended structure; 3_{10}, 3_{10} helical turn; t, turn. [b] The alignment score with the SH3 domain of chicken brain α-spectrin as evaluated by DALI (Holm & Sander, 1993). A higher number indicates a better alignment on a scale from 0 to 9. PsaE has 13% amino acid identity with α-spectrin. This extent of homology places the relationship in the twilight zone (Doolittle, 1990). [c] The primary structure of the SH3 domain of chicken brain α-spectrin (Musacchio et al., 1992) aligned against PsaE with the program DALI (Holm & Sander, 1993). The secondary structure designations (below) are taken from Noble et al. (1993). [d] The primary structure of the SH3 domain of chicken c-Src; the first residue is Thr-85. The boldface symbols indicate the residues whose chemical shift is affected by the binding of a proline-rich peptide (Yu et al., 1992a). Secondary structure assignments (below) are taken from the same work.

at the N-terminus of FNR (Schluchter & Bryant, 1992). The architecture of cyanobacterial PsaE, comprised of a stable SH3 domain and appended loops, suggests that similar structural information might be accessible for PsaE from higher plants and that an explanation for this significant functional difference might be obtainable. A comparison of primary structures indicates that the major differences are found in the N-terminal extension and in the loops and that the SH3 fold should be maintained.

Several studies suggest that PsaE influences the reduction of soluble electron acceptors such as flavodoxin or ferredoxin or both. Weber and Strotmann (1993) found that the rate of ferredoxin-mediated $NADP^+$ photoreduction declined for spinach PS I preparations when PsaE was extracted with chaotropic agents. Addition of recombinant PsaE from *Synechococcus* sp. PCC 7002 could restore about one-half of the lost rate. Flavodoxin reduction was stimulated up to a factor of 2 by the addition of PsaE to reconstituted PS I reaction

center preparations [see Golbeck (1993)]. Both Sonoike et al. (1993) and Rousseau et al. (1993) observed that the presence of PsaE is correlated with faster rates of ferredoxin reduction by PS I complexes, although the latter study found that the yield of reduced ferredoxin was not much affected by PsaE depletion. Attempts to demonstrate cross-linking of ferredoxin and flavodoxin to PsaE were unsuccessful; however, the yields of a cross-linking product between PsaE and PsaF, as well as the yields and patterns of cross-linking of subunits PsaC and PsaD to flavodoxin and ferredoxin, were altered when PsaE was missing (U. Mühlenhoff and D. A. Bryant, unpublished results). While the organization of the acceptor-side proteins seems modified by the absence of PsaE, it is unlikely that the effects noted above describe the primary function of the PsaE subunit, since mutants that contain no PsaE are capable of wild-type growth rates under saturating light conditions (Bryant et al., 1990; Chitnis et al., 1989; Golbeck, 1994; Zhao et al., 1993). The correct placement of the PsaE subunit with respect to the other stromal side proteins and with respect to the transmembrane helices will allow for a better understanding of this role of PsaE, and the coordinates obtained in this study will be useful in locating the protein in the electron density map of the whole complex.

Recent studies with several mutant strains of *Synechococcus* sp. PCC 7002 have shown that mutants devoid of PsaE lack methyl viologen-sensitive, cyclic electron transport around PS I (Yu et al., 1992b, 1993b). Since PsaE binds no known electron-transport cofactors, the mechanism by which PsaE participates in this process is not understood. Studies with thylakoids of *Synechococcus* sp. strain PCC 7002 suggest that PsaE can be cross-linked to a membrane-intrinsic protein that is not a component of the PS I reaction center complex (U. Mühlenhoff and D. A. Bryant, unpublished results). Thus, it is possible that PsaE modulates the binding of some minor electron-transport component to the PS I complex and participates indirectly in cyclic electron transport. Particularly interesting in this context is the role of SH3 domains as mediators of protein–protein interactions in cellular signaling cascades. A similar role for PsaE in facilitating protein interaction during electron transport could be supposed. Further studies, modeled after this working hypothesis, will be required to reveal the precise role of this subunit in PS I-mediated cyclic electron transport.

CONCLUSIONS

The solution structure of PsaE, a small, basic-p*I* water-soluble protein which is associated with the stromal (acceptor or reducing) side of the PS I complex, has been determined. The antiparallel five-stranded β-sheet is well-defined. If one excludes the loops, in particular, the 14-residue CD loop that is a striking feature of the cyanobacterial protein, PsaE shows a remarkable similarity to SH3 domains. These are proteins involved in signal transduction pathways and implicated in protein–protein interactions near membrane surfaces, a function that is likely in PsaE. However, there is little sequence identity with known SH3 domains, and PsaE may be a functionally distinct representative of the SH3 family.

ACKNOWLEDGMENT

The authors are particularly grateful for the gracious help of Drs. Liisa Holm and Gerrit Vriend in the laboratory of Dr. Chris Sander. The homology between PsaE and SH3 domains was discovered by subjecting the PsaE coordinates to their three-dimensional matching programs WHAT IF and DALI using the up-to-date coordinate data bank of the EMBL facility in Heidelberg. Thanks are also due to Georg Tuparev for advice on WHAT IF. Dr. George Rose's guidance was essential in carrying the project to its current end. J. T. J. Lecomte acknowledges Van Veen's contribution to the computational aspects. Many other co-workers and colleagues are thanked for their special support through these studies.

REFERENCES

Andersen, B., Scheller, H. V., & Møller, B. L. (1992) *FEBS Lett. 311*, 169–173.

Baker, E. N. (1980) *J. Mol. Biol. 141*, 441–484.

Booker, G. W., Gout, I., Downing, A. K., Driscoll, P. C., Boyd, J., Waterfield, M. D., & Campbell, I. D. (1993) *Cell 73*, 813–822.

Böttcher, B., Gräber, P., & Boekema, E. J. (1992) *Biochim. Biophys. Acta 1100*, 125–136.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.

Brünger, A. T. (1992) in *X-PLOR. Version 3.1. A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.

Bryant, D. A. (1992) in *The Photosystems: Structure, Function and Molecular Biology* (Barber, J., Ed.) Vol. 11, pp 501–549, Elsevier, Amsterdam, The Netherlands.

Bryant, D. A., Rhiel, E., de Lorimier, R., Zhou, J., Stirewalt, V. L., Gasparich, G. E., Dubbs, J. M., & Snyder, W. (1990) in *Current Research in Photosynthesis* (Baltscheffsky, M., Ed.) Vol. 2, pp 1–9, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Chary, K. V. R., Otting, G., & Wüthrich, K. (1991) *J. Magn. Reson. 93*, 218–224.

Chitnis, P. R., & Nelson, N. (1991) in *Cell Culture and Somatic Cell Genetics of Plants* (Bogorad, L., & Vasil, I. K., Eds.) Vol. 7B, pp 177–224, Academic Press, New York.

Chitnis, P. R., Reilly, P. A., Miedel, M. C., & Nelson, N. (1989) *J. Biol. Chem. 264*, 18374–18380.

Cicchetti, P., Mayer, B. J., Thiel, G., & Baltimore, D. (1992) *Science 257*, 803–806.

Clore, G. M., & Gronenborn, A. M. (1989) *Crit. Rev. Biochem. Mol. Biol. 24*, 479–564.

Clore, G. M., Robien, M. A., & Gronenborn, A. M. (1993) *J. Mol. Biol. 231*, 82–102.

Doolittle, R. F. (1990) *Methods Enzymol. 183*, 99–110.

Falzone, C. J., Kao, Y.-H., Zhao, J., MacLaughlin, K. L., Bryant, D. A., & Lecomte, J. T. J. (1994) *Biochemistry*, (preceding paper in this issue).

Fry, M. J., Panayotou, G., Booker, G. W., & Waterfield, M. D. (1993) *Protein Sci. 2*, 1785–1797.

Golbeck, J. H. (1992) *Annu. Rev. Plant Physiol. Plant Mol. Biol. 43*, 293–324.

Golbeck, J. H. (1993) in *Current Research in Photosynthesis* (Murata, N., Ed.) Vol. 1, pp 487–496, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Golbeck, J. H. (1994) in *The Molecular Biology of Cyanobacteria* (Bryant, D. A., Ed.) Kluwer Academic Publishers, Dordrecht, The Netherlands (in press).

Golbeck, J. H., & Bryant, D. A. (1991) in *Light Driven Reactions in Bioenergetics* (Lee, C. P., Ed.) Vol. 16, pp 83–177, Academic Press, New York.

Havel, T., & Wüthrich, K. (1984) *Bull. Math. Biol. 46*, 673–698.

Havel, T. F., Kuntz, I. D., & Crippen, G. M. (1983) *Bull. Math. Biol. 45*, 665–720.

Holm, L., & Sander, C. (1993) *J. Mol. Biol. 233*, 123–138.

Koch, C. A., Anderson, D., Moran, M. F., Ellis, C., & Pawson, T. (1991) *Science 252*, 668–674.

Kohda, D., Hatanaka, H., Odaka, M., Mandiyan, V., Ullrich, A., Schlessinger, J., & Inagaki, F. (1993) *Cell 72*, 953–960.

Krauss, N., Hinrichs, W., Witt, I., Fromme, P., Pritzkow, W., Dauter, Z., Betzel, C., Wilson, K. S., Witt, H. T., & Saenger, W. (1993) *Nature 361*, 326–331.

Kruip, J., Boekema, E. J., Bald, D., Boonstra, A. F., & Rögner, M. (1993) *J. Biol. Chem. 268*, 23353–23360.

Kuszewski, J., Nilges, M., & Brünger, A. T. (1992) *J. Biomol. NMR 2*, 33–56.

Lagoutte, B., & Vallon, O. (1992) *Eur. J. Biochem. 205*, 1175–1185.

Leszczynski, J. F., & Rose, G. D. (1986) *Science 234*, 849–855.

Li, N., Zhao, J. D., Warren, P. V., Warden, J. T., Bryant, D. A., & Golbeck, J. H. (1991) *Biochemistry 30*, 7863–7872.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Matthews, D. A., Smith, S. L., Baccanari, D. P., Burchall, J. J., Oatley, S. J., & Kraut, J. (1986) *Biochemistry 25*, 4194–4204.

Murzin, A. G., & Chothia, C. (1992) *Curr. Opin. Struct. Biol. 2*, 895–903.

Musacchio, A., Noble, M. E. M., Pautpit, R., Wierenga, R. K., & Saraste, M. (1992) *Nature 359*, 851–855.

Neri, D., Billeter, M., & Wüthrich, K. (1992) *J. Mol. Biol. 223*, 743–767.

Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317–324.

Nilges, M., Kuszewski, J., & Brünger, A. T. (1991) in *Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy* (Hoch, J. C., Poulsen, F. M., & Redfield, C., Eds.) Vol. 225, pp 451–455, Plenum Press, New York.

Noble, M. E. M., Musacchio, A., Saraste, M., Courtneidge, S. A., & Wierenga, R. K. (1993) *EMBO J. 12*, 2617–1624.

Parrett, K. P., Mehari, T., & Golbeck, J. H. (1990) *Biochim. Biophys. Acta 1015*, 341–352.

Powers, R., Garrett, D. S., March, C. J., Frieden, E. A., Gronenborn, A. M., & Clore, G. M. (1993) *Biochemistry 32*, 6744–6762.

Rousseau, F., Sétif, P., & Lagoutte, B. (1993) *EMBO J. 12*, 1755–1765.

Sander, C., & Schneider, R. (1991) *Proteins: Struct., Funct., Genet. 9*, 56–68.

Schluchter, W. M., & Bryant, D. A. (1992) *Biochemistry 31*, 3092–3102.

Sixma, T. K., Pronk, S. E., Kalk, K. H., Wartna, E. S., van Zantem, B. A. M., Witholt, B., & Hol, W. G. J. (1991) *Nature 351*, 371–378.

Sixma, T. K., Stein, P. E., Hol, W. G. J., & Read, R. J. (1993) *Biochemistry 32*, 191–198.

Sonoike, K., Hatanaka, H., & Katoh, S. (1993) *Biochim. Biophys. Acta 1141*, 52–57.

Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L., & Read, R. J. (1992) *Nature 355*, 748–750.

Vriend, G. (1990) *J. Mol. Graphics 8*, 52–56.

Wagner, G., Braun, W., Havel, T. S., Schaumann, T., Go, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611–639.

Weber, N., & Strotmann, H. (1993) *Biochim. Biophys. Acta 1143*, 204–210.

Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J., & Matthews, B. W. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9257–9261.

Wynn, R. M., Omaha, J., & Malkin, R. (1989) *Biochemistry 28*, 5554–5560.

Yu, H., Rosen, M. K., Shin, T. B., Seidel-Dugan, C., Brugge, J. S., & Schreiber, S. L. (1992a) *Science 258*, 1665–1668.

Yu, L., Golbeck, J. H., Zhao, J. D., Schluchter, W., Mühlenhoff, U., & Bryant, D. (1992b) in *Current Research in Photosynthesis* (Murata, N., Ed.) Vol. 1, pp 565–568, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Yu, L., Zhao, J., Lu, W., Bryant, D. A., & Golbeck, J. H. (1993a) *Biochemistry 32*, 8251–8258.

Yu, L., Zhao, J., Mühlenhoff, U., Bryant, D. A., & Golbeck, J. H. (1993b) *Plant Physiol. 103*, 171–180.

Zhao, J. D., Li, N., Warren, P. V., Golbeck, J. H., & Bryant, D. A. (1992) *Biochemistry 31*, 5093–5099.

Zhao, J., Snyder, W. B., Mühlenhoff, U., Rhiel, E., Warren, P. V., Golbeck, J. H., & Bryant, D. A. (1993) *Mol. Microbiol. 9*, 183–194.